Discussion 5.1 – Model Ensemble Design, Feature Scaling, and Regularization Strategies

The following discussion questions are designed to encourage critical thinking and deeper exploration of ensemble modeling, feature engineering, and regularization in predictive systems.

1. **Adaptive Weighting vs. Simple Averaging**  
   In ensemble learning, why might a data-driven weighting method, such as Ridge regression, outperform a simple average of predictions? Consider scenarios involving regime shifts and varying signal-to-noise ratios across models.
2. **Feature Scaling Considerations**  
   When combining features that are all measured on the same scale (e.g., dollar values), is feature standardization still necessary? How would your answer change if the feature set later included heterogeneous measures such as ratios, percentages, and categorical encodings?
3. **Interpreting High Training R² in Ensembles**  
   Why might training-set R² be unusually high when the features themselves are outputs from other models? How would you design a backtesting or cross-validation procedure to obtain a more realistic measure of generalization performance?
4. **Choosing Between Ridge, Lasso, and Elastic Net**  
   In what situations might L1-based methods (Lasso) or a hybrid approach (Elastic Net) be preferable to Ridge regression for weighting ensemble members? Discuss the trade-offs in terms of sparsity, interpretability, and stability under different data regimes.

**Required:** Read all your peers' posts, then comment meaningfully on two or more.